

Key

Chapter 5 Review Packet- Intro to Statistics

MULTIPLE CHOICE:

Sale of eggs that are contaminated with salmonella can cause food poisoning among consumers. A large egg producer takes an SRS of 200 eggs from all the eggs shipped in one day. The laboratory reports that 9 of these eggs had salmonella contamination. Unknown to the producer, 0.1% of all eggs shipped had salmonella. Use this for the next four questions.

66. In this situation,

- A
- (a) 0.1% is a parameter and 9 is a statistic.
 - (b) 9 is a parameter and 0.1% is a statistic.
 - (c) both 0.1% and 9 are parameters.
 - (d) both 0.1% and 9 are statistics.
 - (e) 0.1% is an estimate and 9 is a margin of error.

$$9/200 = 4.5\%$$

67. Based on the sample data, the producer estimates that the proportion of contaminated eggs in the population is about

- D
- (a) 0.2%
 - (b) 0.045%
 - (c) 3%
 - (d) 4.5%
 - (e) 20%

68. A statistician tells the producer that the margin of error for a 95% confidence statement for these data is about plus or minus 3 percentage points. What is the confidence statement for this information?

- B
- (a) We are 95% sure that the eggs are good
 - (b) We are 95% confident that the true percent of eggs that are contaminated is between 1.5% and 7.5%.
 - (c) We are 95% confident that the percent of eggs in the sample that are contaminated is between 1.5% and 7.5%.
 - (d) We cannot make a statement from the information given

69. If the producer took an SRS of 400 eggs instead of 200, the new margin of error would be

- B
- (a) the same as before, because the population of eggs is the same.
 - (b) smaller than before, because the sample is larger.
 - (c) larger than before, because the sample is larger.
 - (d) random in size, could be either larger or smaller than before.
 - (e) can't tell, because the size of the sample doesn't control the margin of error.

84. A radio talk show invites listeners to call a telephone number to vote "Yes" or "No" on whether they support a bond issue for a new school. About 1500 people call in. Over 80% say "No." As an estimate of community opinion, this result is

- E
- (a) accurate to within $\pm 3\%$ with 95% confidence.
 - (b) not trustworthy because of nonsampling errors.
 - (c) not precise because the sample size 1500 is too small.
 - (d) unethical due to lack of informed consent.
 - (e) badly biased due to voluntary response.

85. You must choose a simple random sample of 7 packages from a shipment of 80 packages of vaccine for testing. You label the packages 01, 02, 03, ..., 80 and then use the random number table to select your sample. The chance that any one package is chosen for the sample is

- C
- (a) random because the sample is random.
 - (b) can't say without looking at the random number table.
 - (c) 1 in 80.
 - (d) 7 in 80.
 - (e) 7 in 100.

86. You will take a simple random sample of 3 cases from a population of 11 cases of wine, which are marked A, B, C, D, E, F, G, H, I, J, K. You assign numbers as labels to the cases in order to use the random digit table. Which of the following is an incorrect assignment?

	A	B	C	D	E	F	G	H	I	J	K
(a)	01	02	03	04	05	06	07	08	09	10	11
(b)	00	01	02	03	04	05	06	07	08	09	10
(c)	1	2	3	4	5	6	7	8	9	10	11
(d)	0	1	2	3	4	5	6	7	8	9	10
(e)	Both (c) and (d) are incorrect.										

88. Which of the following is a source of *nonsampling error* in a sample survey?

- (a) Voluntary response sampling.
- (b) Using voter registration lists as the sampling frame.
- (c) Some subjects do not tell the truth.
- (d) Both (b) and (c).
- (e) None of these.

90. An opinion poll asks a sample of 1100 people whether they support reducing the number of legal immigrants to the U.S.; 53% of these 1100 people say "Yes." The number 53% is a

- (a) margin of error.
- (b) statistic.
- (c) bias.
- (d) parameter.
- (e) reliability.

96. Which of these statements about a random number table is true?

- (a) Each line contains exactly the same number of 0's and 1's.
- (b) The chance that any pair of adjacent digits is 88 is $1/100$.
- (c) The chance that any pair of adjacent digits is 89 is $1/99$.
- (d) No row of the table can consist of all 9's.

— there are 100 pairs
dic
00-9

97. Which of the following is not a source of nonsampling error:

- (a) inability to contact a subject.
- (b) accidentally throwing away the completed survey questions.
- (c) choosing the sample from shoppers at a suburban mall.
- (d) poorly worded questions.

98. Which of the following is not true of a simple random sample of size 1000 chosen from a population of size 4,000,000?

- (a) Every individual in the population has the same chance of selection as every other individual.
- (b) Every set of 1000 individuals has the same chance of being the sample as every other set of 1000 individuals.
- (c) Every individual of the population has chance 1 in 1,000 of being included in the sample.
- (d) Every set of 500 individuals has the same chance of being included in the sample as every other set of 500 individuals.

104. The United Presbyterian Church recently took a sample of opinion in the church. The overall sample "contains independent random samples of 1537 members, 1400 elders, 1513 pastors and 714 other clergy." This sampling design is a

- (a) multistage sample.
- (b) voluntary response sample.
- (c) simple random sample.
- (d) stratified sample.
- (e) None of the above

106. Some common sources of nonsampling error in samples of human populations are

- (a) using voluntary response samples; some subjects lie.
- (b) some subjects lie; some subjects can't be contacted.
- (c) some subjects can't be contacted; drawing a sample from names in a telephone directory.
- (d) Both (b) and (c).
- (e) All of (a), (b), and (c).

The General Social Survey (GSS), conducted by the National Opinion Research Center at the University of Chicago, is a major source of data on social attitudes in the U.S. Once each year 1500 adults are interviewed in their homes all across the country. The subjects are asked their opinions about sex and marriage, attitudes toward women, welfare, foreign policy and many other issues. **The next few questions** concern the GSS.

111. The population for the GSS is

- A ☒ (a) all adult residents of the U.S. (b) the University of Chicago.
(c) the 1500 persons interviewed. (d) the list of questions asked.

112. The GSS finds that 28% of the 1500 people interviewed do not approve of capital punishment. The number 28% is

- D (a) a confidence level. (b) a random digit. (c) a parameter. ☒ (d) a statistic.

113. The GSS begins by dividing the 3000 counties in the country up into urban, rural, and suburban; then a separate sample is chosen at random from each group. This is a

- B (a) simple random sample. ☒ (b) stratified random sample.
(c) systematic random sample. (d) voluntary response sample.

115. If the GSS interviewed 1500 adults from Illinois (population 12,000,000) rather than 1500 from the entire U.S. (population 280,000,000) the variability of Illinois would be

- C (a) smaller because Illinois has a smaller population.
(b) larger because Illinois has a smaller population.
☒ (c) the same because population size doesn't matter.
(d) either smaller or larger because it's random.

116. A potential source of nonsampling error in the GSS is

- A ☒ (a) some subjects refuse to give their opinions.
☒ (b) the sample design chooses dwelling units, so homeless people are left out. *sampling error*
☒ (c) both blacks and whites are included in the sample.
(d) Both (a) and (b), but not (c).

117. Increasing the sample size of an opinion poll will

- B ☒ (a) reduce the bias of the poll result.
☒ (b) reduce the variability of the poll result.
(c) not affect the bias or the variability
(d) all of the above.
(e) (a) and (b) but not (c).

118. A member of Congress receives 1128 letters about proposed legislation that would provide government insurance for nursing-home care. Over 80% oppose the legislation. This sample can't be trusted because

- C ☒ (a) it comes from a simple random sample.
☒ (b) it includes people with no interest in the issue.
☒ (c) there is bias due to voluntary response.
(d) it comes from a stratified random sample.

119. A sample of student opinion at a Big Ten university selects an SRS of 200 of the 30,000 undergraduate students and a separate SRS of 100 of the 5,000 graduate students. This kind of sample is called a

- B (a) simple random sample. ☒ (b) stratified random sample.
(c) systematic random sample. (d) multistage sample.

121. A quality control inspector on an assembly line making microwave ovens randomly chooses one of the first ten ovens manufactured each day. This oven and every tenth oven thereafter gets inspected. This is called

- D (a) voluntary response sampling.
(b) an experiment.
(c) simple random sampling.
☒ (d) systematic random sampling.
(e) stratified random sampling.

124. Voluntary response polls almost always suffer from

- (a) the placebo effect.
- ☒ (b) high bias.
- (c) highly variable results.
- (d) randomization.

Professor Lilli Gans wishes to study the relationship between a person's zodiac sign and his or her political opinions. She obtains the birthdays of all 816 students enrolled in her Astrology 101 course. Then she uses a random digits table to choose 10 students from each of the twelve zodiac signs. For example, students born from March 21 until April 19 have the Aries zodiac sign, and Professor Gans chooses 10 Aries students at random. She uses a different part of the table of random digits to choose students having each sign. After obtaining her sample, Professor Gans has the selected students fill out a questionnaire. **The next two questions** concern this study.

125. Professor Gans has obtained

- (a) A simple random sample from her class.
- (b) A systematic random sample from her class.
- (c) A stratified random sample, where the strata are students with different political opinions.
- ☒ (d) A stratified random sample, where the strata are students with different zodiac signs.

126. One of the questions in Professor Gans' questionnaire asks whether the student has ever given money to the American Nazi Party. Any students who have done so are likely to lie and say "No," rather than admit the truth. This is an example of

- (a) confounding
- (b) voluntary response.
- (c) sampling error
- ☒ (d) nonsampling error

128. If we took a SRS of 1700 people from California (population 32 million) and a SRS of 1000 people from Detroit (population 1 million) which sample would have the smaller variability? \rightarrow bigger sample

- ☒ (a) Detroit, because $1000/1,000,000$ is more than $1700/32,000,000$.
- (b) Detroit, because it has the smaller population.
- ☒ (c) California, because the sample size (1700) is larger than for Detroit (1000).
- (d) California, because it has the larger population.
- (e) Both would be the same, because simple random samples (SRS) are taken in both places.

132. A sample survey to study the effect of race on political affiliation takes a sample consisting of 500 randomly selected whites, 500 randomly selected blacks, and 500 randomly selected Hispanics. This is an example of a

- (a) voluntary response sample.
- (b) simple random sample.
- (c) systematic random sample.
- ☒ (d) stratified random sample.
- (e) None of the above.

135. Which of the following are sources of nonsampling errors?

- ☒ (a) You can't ever contact some of the students in your sample.
- ☒ (b) Some of the students you contact refuse to answer.
- ☒ (c) You choose students by walking through the crowd at a football game.
- ☒ (d) (a) and (b) but not (c).
- (e) All of (a), (b), and (c).

137. A *simple random sample* is

- (a) any sample selected by using chance.
- (b) any sample that gives every individual the same chance to be selected.
- ☒ (c) a sample that gives every possible sample of the same size the same chance to be selected.
- ☒ (d) a sample that selects equal numbers of individuals from each stratum.
- ☒ (e) a sample that contains the same percent of each subgroup in the population.

FREE RESPONSE

- Take an SRS of 5 from the following sampling frame. Start at line 110 in the Table. Start your labels at 0 label down the columns.

00 Matthew	05 Nicole	10 Greg	15 Daniel	20 Richard
01 Francis	06 Sarah	11 Nicholas	16 Karen	21 Megan
02 Tara	07 Leah	12 Jennifer	17 Michael	22 Ryan
03 Jason	08 Joseph	13 Wendy	18 Emily	23 Danielle
04 Chris	09 Terry	14 Andy	19 John	24 Nelson

- The following sampling frame contains both males and females. The males are in **BOLD**. Take a stratified sample of 6 males and 3 females. For the males, use line 105 of the table. For females, use line 131 of the table. Be sure to show your labels below, so I know where you started and how you labeled.

01 Jimmy	05 Larry	07 Jake	09 Sue	14 Erik	18 Cliff
02 Roy	03 Sarah	07 Gretchen	10 Gwen	15 Felix	19 Derek
01 Nicole	04 Annie	08 Kelly	11 Bob	12 Colleen	20 Danny
03 Mark	05 Patty	08 John	11 Marge	13 Lisa	21 Rich
02 Kate	06 Alex	09 Colin	12 Kevin	16 Steve	14 Amy
04 Thomas	06 Jason	10 Chris	13 Bart	17 Ian	22 Mitch

Males: 07, 19, 14, 17, 13, 15

Females: 05, 04, 07

Jake, Derek, Erik, Ian, Bart, Felix

Patty, Annie, Gretchen

- Using the list from the previous problem (ignore gender this time), take a systematic random sample by doing the following:
 - Assign each person a number, starting at 1
 - Randomly select a number between 1 and 10 (you can use your calculator for this). This person will be the first person in your sample.
 - Now every 6th person from the one you selected (example: person 3, 9, 15, 21, etc.). This is your sample.

01 Jimmy	07 Larry	13 Jake	19 Sue	25 Erik	31 Cliff
02 Roy	08 Sarah	14 Gretchen	20 Gwen	26 Felix	32 Derek
03 Nicole	09 Annie	15 Kelly	21 Bob	27 Colleen	33 Danny
04 Mark	10 Patty	16 John	22 Marge	28 Lisa	34 Rich
05 Kate	11 Alex	17 Colin	23 Kevin	29 Steve	35 Amy
06 Thomas	12 Jason	18 Chris	24 Bart	30 Ian	36 Mitch

randInt(1, 10) → 04

Mark, Patty, John, Marge, Lisa, Rich

- What two methods of sampling are always biased?

voluntary response, convenience sampling

- What is Bias?

systematically favoring one response

- What is variability?

spread of a sampling distribution

- What is the difference between bias and variability?

accuracy

precision

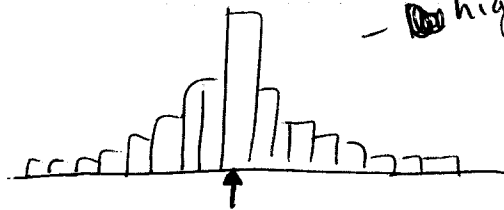
8. Identify the following pictures as high or low bias, and high or low variability:

a)



- high bias
- high var.

c)



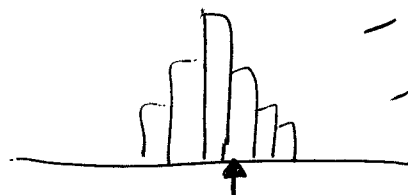
- low bias
- high var.

b)



- high bias
- low var.

d)



- low bias
- low var.

9. In order to determine the true feelings of community members about a law mandating that bike riders wear helmets, a local municipality randomly selects 100 people from local tax returns to sample. They then call each household and ask for the head of the household. Once that person is on the phone, they ask their question "Are you in favor of life-saving law that mandates people wear bike helmets?" They find that 82% of them are in favor of the law.

- What is the sample? 100 people from community
- What is the sampling frame? tax returns
- What is the population? all community members
- What is the parameter of interest? % of community members in favor of law
- Is 82% a statistic or a parameter? statistic
- List all possible sources of bias/error in this sampling method. Then identify each as sampling or non-sampling error.

- samp. • undercoverage - only use those people who do tax ret
- samp • undercoverage - only those w/ phones
- non-samp • nonresponse - people not home or hang up when call
- wording of question - slanted in favor of law

10. Identify the numbers in the following problems as statistics or parameters:

- A supermarket takes a random sample of 150 of its shoppers carts and finds that 6% of them buy organic produce. This is interesting since the National Average is 10%.
6% - statistic
10% - parameter
- A scientist takes a random sample of 50 pitches from a starting Phillies pitcher and finds an average speed of 85 mph. However the Phillies pitching coach says that the pitcher has always averaged 88 mph.
85 mph - statistic
88 mph - parameter
- John decides to flip a coin. He flips a quarter 90 times and finds that it lands on heads 55% of the time. He then decides to do the same for a penny. He finds the penny lands on heads 49% of the time.
55% - statistic
49% - statistic

Sample
↓

11. Jimmy wants to take a good sample of students in his high school. He does an SRS of 150 students. His high school has 1200 people in it. What is wrong with his sample?

the population (1200) is not greater than 10n.
10(150)

12. A newspaper does a survey of 1245 people and finds that 254 of them do not have health care. The margin of error is 4%. They do this sample by standing outside of a hospital and sampling every 5th person that comes through the door.

- a. What is the population? All people
- b. What is the sample? 1245 people
- c. What is the sampling frame? people who enter hospitals
- d. What is the statistic in this problem? $254/1245 = 20.4\%$
- e. What is the parameter ^{of} interest? the true % of people who don't have health care
- f. Create a confidence statement from this information
- g. This sampling method is most certainly biased. However, is the TRUE percent of people who do not have health care lower or higher than the percent found in this sample? Why??
higher. People who don't have healthcare don't come to hospitals as much
- h. How could I reduce the margin of error? take a bigger sample
- i. List all possible sources of bias/error and then note for each if it is a sampling or non-sampling error.

- samp. • undercoverage - only sampled people who went to hospital
- samp. • systematic (not random) sample - every 5th person
- non-samp. • non response - some people might not answer

13. Identify the following as sampling errors or non-sampling errors, and also give it a more specific name (vocab!)

- a. A respondent lies to a surveyor nonsampling - response error
- b. A respondent is confused by the way a question is phrased (even though the question is fairly clear) nonsampling - response error
- c. A researcher records a piece of data wrong non sampling - processing error
- d. A researcher asks people to call in to express their opinions Sampling - voluntary response error
- e. A researcher states that 70% of all traffic accidents are caused by drunk drivers. He then asks people what they think about toughening the drunk driving laws. Sampling - wording of questions
- f. Some households refuse to send back questionnaire that was sent to them. nonsampling - non response